REPORT DOCUMENTATION PAGE		Form Approved OMB No. 0704-0188
1. AGENCY USE ONLY (leave blank)	2. REPORT DATE 15-Jul-96	3. REPORT TYPE AND DATES COVERED Progress Report 1 Nov 94-31 Dec 95
4. TITLE AND SUBTITLE Research problems in wireless communication networks		5. FUNDING NUMBERS F 49620-95-1-0061
6. AUTHOR(S) Leandros Tassiulas		AFOSR-TR-96 0447
7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES) Polytechnic University 6 Metrotech Center Brooklyn, NY 11201		
9. SPONSORING/MONITORING AGENCY NAME(S) AND ADDRESS(ES) Air Force Office of Scientific Research 110 Duncan Avenue, Suite B115 Bolling AFB, DC 20332		10. SPONSORING/MONITORING AGENCY REPORT NUMBER
11. SUPPLEMENTARY NOTES		
12a. DISTRIBUTION/AVAILABILITY STATEMENT DISTRIBUTION STATEMENT A Approved for public release Distribution Unlimited		12b. DISTRIBUTION CODE

13. ABSTRACT (maximum 200 words)

The work performed under the support of the aforementioned AFOSR grant includes research on terrestrial wireless access systems and satellite networks. In a terrestrial network, a multilayer cellular architecture was investigated. It consists of layers with dense frequency reuse for mobiles close to base stations and with sparser frequency reuse for remote mobiles. Channel control schemes that increase the spectrum utilization were proposed and studied. In the digital wireless systems of the near future, a mobile terminal will have control over the transmission power, channel selection and base station assignment. Control algorithms that compute and assign all those quantities in order to increase the efficiency and robustness of the system were studied. The optimal solutions in the two base station case were obtained, while heuristic algorithms for larger systems were proposed and evaluated. It appears that the joint control of power, channel and base station assignment can significantly increase the performance of the system. The problem of management and control in networks with time-varying topology as it arises in low earth orbit satellite networks was studied. Control policies that use predictions of the network topology evolution were proposed and studied. The magnitude of improvement in the performance using predictive policies was significant.

14. SUBJECT TERMS			15. NUMBER OF PAGES
Wireless network management, channel control, power control, satellite networks			33
			16. PRICE CODE
17. SECURITY CLASSIFICATION OF REPORT Unclassified	18. SECURITY CLASSIFICATION OF THIS PAGE Unclassified	19. SECURITY CLASSIFICATION OF ABSTRACT Unclassified	20. LIMITATION OF ABSTRACT

Standard Form 298 (Rev 2-89)

FINAL TECHNICAL REPORT

- **AFOSR Contract Number:** #F49620-95-1-0061
- **Date:** 15 July, 1996
- **Period Covered By This Report:** 1 November 1994- 31 December 1995
- **Title:** "Research problems in wireless communication networks"
- **PI Name:** Leandros Tassiulas
- **PI Institution:** Dept. of Electrical Engineering, University of Maryland, College Park, MD 20742 (most of the work reported here has been done while the PI was with Polytechnic University, Brooklyn, NY 11201)
- **Submitted to:**
  Air Force Office of Scientific Research
  110 Duncan Avenue, Suite B115,
  Bolling AFB, DC 20332
- **Number of Pages:** 33

UNCLASSIFIED
Distribution Unlimited

Executive Summary

Abstract

The work performed under the aforementioned AFOSR grant includes research on terrestrial wireless access systems and satellite networks. In the digital wireless systems of the near future, a mobile terminal will have control over the transmission power, channel selection and base station assignment. Control algorithms that compute and assign all those quantities in order to increase the efficiency and robustness of the system were studied. The optimal solutions in the two base station case were obtained [7], while heuristic algorithms for larger systems were proposed and evaluated [6]. It appears that the joint control of power channel and base station assignment can significantly increase the performance of the system. The problem of management and control in networks with time-varying topology as it arises in low earth orbit satellite networks was studied [11]. Control policies that use predictions of the network topology evolution were proposed and evaluated. The magnitude of improvement in the performance using predictive policies was significant. The main results and algorithms resulted from the above effort are summarized here.

1 Joint power, channel and base assignment for wireless access

The provision of Personal Communication Services (PCS) is the goal of the evolution of integrated communication systems. The basic PCS philosophy is that the underlying telecommunication infrastructure will provide user-to-user, location independent, communication services. The services supported by PCS generate a large volume of communication traffic with diverse burstiness characteristics and quality of service requirements. The wireless network should be able to support this traffic and to meet the quality of service requirements. Efficient utilization of the limited radio spectrum will be vital for meeting the anticipated traffic demands of PCS.

In the digital wireless systems of the near future, there will be great flexibility in spectrum management and control. All the channels will be potentially available to all base stations, to be allocated in a dynamic fashion. Also the transmission powers both of the base stations and the users will be controllable. This great flexibility will increase considerably the traffic capacity of the system if it is managed properly.

In a network with dynamic power and channel control, the channels the powers and the base stations can be reassigned at any time, even in the middle of a call. As the mobiles change positions, the powers of the received signals change and a reassignment (hand-off) might be required to retain the connection. When a new call request arises, a reassignment of the existing calls might be required to accommodate it. The fundamental problem underlying any phase (hand-off, new connection, etc.) of a dynamic resource allocation algorithm in a wireless network is the following.

P: Given a number of channels, a number of mobiles, a number of base stations and the path losses among them, assign transmission powers, channels and base stations such that all mobiles have a connection.

The above problem is studied and a solution is obtained for the case of two base stations [7]. It is shown here that the optimal assignment is obtained by the solution of the maximum matching problem in an appropriate graph. The graph has as nodes the mobiles and links connect any two mobiles which can share the same channel by appropriate base station assignment and power selection. There is one-to-one correspondence between matchings and frequency assignments where the pairs of the mobiles that correspond to the links of the matching share the same channel. A maximum matching corresponds to maximal channel reuse or equivalently to provision of connections using the minimal number of channels.

Based on the optimal assignment algorithm the traffic capacity, appropriately defined, is studied as well. Let $C_N^f(C_N^r)$ be the minimum number of channels needed to establish a forward (reverse) connection to N mobiles, randomly and uniformly distributed, when the transmission power is fixed. The traffic capacity T_N^f in the forward direction without power control is defined as

$$T_N^f = \frac{N}{E[C_N^f]}, \tag{1}$$

and similarly for the reverse direction. For both the forward and reverse channels T_2^f and T_2^r are computed, while in general, T_N^f and T_N^r are obtained by simulation. For the forward channel the limit of the forward traffic capacity T_N^f, as N increases, is obtained analytically. It turns out that for small N, $T_N^f > T_N^r$ while for large N the inequality is reversed. In general the capacity is different for the forward and reverse channel. The case of power control is considered next. For every configuration of N mobiles, the minimum number of channels necessary to establish a connection in the forward and reverse direction for every mobile, $\tilde{C}_N^f$ and $\tilde{C}_N^r$ respectively, are equal. Hence the corresponding capacities $\tilde{T}_N^f$ and $\tilde{T}_N^r$, defined as in (1), are equal, $\tilde{T}_N^f = \tilde{T}_N^r = \tilde{T}_N$. Note that by $\tilde{C}$ and $\tilde{T}$ we denote the minimum number of channels needed in the system and the corresponding traffic capacity respectively, when the powers are controllable parameters. We obtain $\tilde{T}_N$ by simulation and compare it with T_N^f and T_N^r. As it is expected, a considerable improvement on the capacity is observed. The problem of two way channel assignment is also studied and similar results are obtained.

1.1 The joint problem

In this section we introduce some notation and define the allocation problem rigorously. Even though we will study exclusively the case of two base stations in this paper, we formulate the general problem. There are L communication channels available in the system. The channels may

be either frequency bands in an FDMA system or a carrier and a time slot in a TDMA system or different codes in a CDMA system. There are M base stations and N mobiles at arbitrary locations with respect to the base stations. The path loss coefficients G_{ij} between any base station i and mobile j are provided. They characterize completely the propagation properties of the system in the sense that when i transmits power P_i, j receives power $G_{ij}P_i$. We denote by P_{il}^f the transmitted power from base station i in the forward channel l. Similarly P_{jl}^r denotes the transmitted power from mobile j in the reverse channel l. Cochannel interference is the prevailing interference type; this is equal to $\sum_{k \neq i} G_{kj} P_{kl}^f$ in mobile j receiving from base station i at channel l. The carrier to interference ratio $(C/I)_{jl}^f$ at mobile j in channel l is equal to

$$(C/I)_{jl}^f = \frac{G_{ij} P_{il}^f}{\sum_{k \neq i} G_{kj} P_{kl}^f}. \tag{2}$$

The interference constraint at mobile j that receives on channel l is satisfied if

$$(C/I)_{jl}^f \geq T, \tag{3}$$

where T is a threshold imposed by physical layer's constraints. The constraint (3) is on the area mean carrier power and the interference power. T is selected such that (3) guarantees that the effect of fast and slow fading will not be detrimental on the link quality. Similarly the carrier to interference ratio $(C/I)_{il}^r$ at base station i in channel l is equal to

$$(C/I)_{il}^r = \frac{G_{ij} P_{jl}^r}{\sum_{n \neq j} G_{in} P_{nl}^r}. \tag{4}$$

The reverse radio link from mobile j to base station i at channel l satisfies the interference constraints if

$$(C/I)_{il}^r \geq T. \tag{5}$$

The problem of joint channel power and base station allocation is illustrated in figures 1 and 2. Clearly the three problems are interrelated. For certain channel allocations and base station assignments there may be power vectors that satisfy the interference constraints while for others may be not. Therefore these problems need to be considered jointly.

1.1.1 Forward channel assignment problem

A forward channel assignment is specified by a function $C^f() : \{1,..,N\} \rightarrow \{1,..,L\}$ with the interpretation that $C^f(j)$ is the channel at which mobile j is receiving. A base station assignment for the forward channel is specified by a function $B^f() : \{1,..,N\} \rightarrow \{1,...,M\}$ where $B^f(j)$ is the base station from which j is receiving. A forward channel assignment $C^f()$ and base station

assignment $B^f()$ are *jointly admissible* if at most one mobile is assigned to each base station in each channel. In other words if

$$C^f(i) = C^f(j) \Rightarrow B^f(i) \neq B^f(j).$$

A *forward channel-base station assignment* is a pair (C^f, B^f) of jointly admissible forward channel and base station assignments respectively. A forward channel-base station assignment is feasible if there exists a transmission power assignment such that at each mobile the carrier to interference ratio from its assigned base station in the assigned channel, exceeds the required threshold T. That is

$$\max_{P^f \geq 0}\{\min_{j=1,..,N}\{\frac{G_{B^f(j)j}P^f_{B^f(j)C^f(j)}}{\sum_{k\neq B^f(j)} G_{kj}P^f_{kC^f(j)}}\}\} \geq T \tag{6}$$

The resource allocation problem can be stated as follows

$$\max_{(C^f,B^f)}\{\max_{P^f \geq 0}\{\min_{j=1,..,N}\{\frac{G_{B^f(j)j}P^f_{B^f(j)C^f(j)}}{\sum_{k\neq B^f(j)} G_{kj}P^f_{kC^f(j)}}\}\}\} \tag{7}$$

Where (C^f, B^f) is a jointly admissible
forward channel base station assignment

If the optimal value of the objective function is larger than the threshold T then the channel and base assignment (C^f, B^f) and the power assignment P^f that achieve the maximum provide a feasible allocation.

1.1.2 Reverse channel assignment problem

A reverse channel assignment is specified by a function $C^r() : \{1, .., N\} \rightarrow \{1, .., L\}$ with the interpretation that $C^r(j)$ is the channel at which mobile j is transmitting. A base station assignment for the reverse channel is specified by a function $B^r() : \{1, .., N\} \rightarrow \{1, ..., M\}$ where $B^r(j)$ is the base station to which j is transmitting. The reverse channel assignment problem can be formulated similarly to the forward channel assignment problem.

1.1.3 Two way channel assignment problem

There are several versions of the two way channel assignment problem depending on the constraints we pose. In the existing analog system a channel is prespecified as forward or reverse and it can be used only in one direction. This is certainly not a physical constraint since in principal the same channel can be used in both directions simultaneously in sufficiently spatially separated locations.

In the future digital systems this constraint can be eliminated and a channel may be used for communication in both directions. Moreover another constraint, usually imposed, that affects the system capacity is that a mobile should communicate with the same base station in both directions. If we ignore the practical constraints imposed in specific systems, each mobile may use any channel and base station for each one of the reverse and forward connections. An *admissible two way channel base station assignment* in this case is a quatraple $\mathbf{CB}() = (C^f(), C^r(), B^f(), B^r())$ for which the following are satisfied

$C^f(i) \neq C^r(i),\ i = 1,..,N$ A mobile cannot talk and listen in the same channel
$C^f(i) = C^f(j) \Rightarrow B^f(i) \neq B^f(j),\ i,j = 1,..,N,\ i \neq j.$ A base station can talk at most to one mobile per channel.
$C^r(i) = C^r(j) \Rightarrow B^r(i) \neq B^r(j),\ i,j = 1,..,N,\ i \neq j.$ A base station can listen at most to one mobile per channel.
$C^f(i) = C^r(j) \Rightarrow B^f(i) \neq B^r(j),\ i,j = 1,..,N,\ i \neq j.$ A base station cannot talk and listen at the same channel.

The problem can be stated as follows

$$\max_{\mathbf{CB}}\{\max_{(P^f,P^r)\geq 0}\{\min_{j=1,..,N}\{\frac{G_{B^f(j)j}P^f_{B^f(j)C^f(j)}}{\sum_{k\neq B^f(j)} G_{kj}P^f_{kC^f(j)}}\},\min_{j=1,..,N}\{\frac{G_{B^r(j)j}P^r_{jC^r(j)}}{\sum_{n\neq j} G_{B^r(j)n}P^r_{nC^r(j)}}\}\}\} \tag{8}$$

Where $\mathbf{CB}() = (C^f, C^r, B^f, B^r)$ *is an admissible two way channel base station assignment*

If the optimal value of the objective function is larger than the threshold T then the channel assignment $\mathbf{CB}$ and the power assignment P^f and P^r that achieves the maximum provide a feasible allocation. The above optimization problems arise constantly in the management and control of a wireless network. In the generality that it is stated these problems are clearly hard optimization problems.

1.2 Optimal assignment

In this section we propose an algorithm that achieves the optimal channel, base station and power assignment, for the case of two base stations. The algorithm minimizes the number of channels needed to establish communication for all mobiles in the forward and reverse directions. Note that minimizing the number of channels is equivalent to the satisfiability problem (P). The procedure we propose is based on the solution of a maximum matching problem in an appropriate graph and is applicable for both the cases of power control (i.e the powers are controllable parameters) and

non power control (i.e. fixed powers). Before proceeding with the description of the algorithm we present some preliminary results on the conditions that should be satisfied such that two transmitter -receiver pairs can make use of the same channel. Since problem (P) is the same in the forward and reverse channel assignment we focus on the forward channel in the following.

Lets denote by A and B two transceivers that may communicate with the transceivers 1 and 2, respectively (fig.3). We consider as forward channel a channel used for communication from transmitter 1(2) to receiver A(B) and as reverse a channel that is used for communication in the opposite direction. The next theorem states the conditions for having channel reuse. Although the results of the following theorem could follow from the more general solution for M base stations and M mobiles presented in [12]

Theorem 1 (a) *Both receivers A and B can receive on the same forward channel iff the powers* P_{1l}^f *and*P_{2l}^f *transmitted by transmitters 1 and 2 satisfy the following relation*

$$\frac{T}{(G_{1A}/G_{2A})} \leq \frac{P_{1l}^f}{P_{2l}^f} \leq \frac{(G_{2B}/G_{1B})}{T}. \tag{9}$$

(b) *When the power is controlled the necessary and sufficient condition for transceivers A and B to receive on the same forward channel and/or transmit on the same reverse channel is the following*

$$\sqrt{\frac{G_{2B}G_{1A}}{G_{1B}G_{2A}}} \geq T, \tag{10}$$

and the transmission powers P_{1l}^f, P_{2l}^f *that achieve it, if possible, are such that*

$$\frac{P_{1l}^f}{P_{2l}^f} = \sqrt{\frac{G_{2B}G_{2A}}{G_{1B}G_{1A}}}. \tag{11}$$

From part (b) of the above theorem we can easily conclude that when the powers are controlled the forward and reverse channel assignment problems are equivalent, in the sense that it is stated in the following corollary.

Corollary 1 *When the powers are adjustable parameters, two transceivers that communicate with another pair of transceivers can share one forward channel iff they can share one reverse channel.*

It should be noted that the results of theorem 1 and corollary 1 are valid under the assumption of no background noise. In this case if relation (10) is satisfied then we can always adjust the transmitted powers according to (11) such that both the constraints are satisfied, even if the maximum allowable transmitted power is finite and limited. If we take into consideration the existence of background

noise then in the case that the maximum allowable transmitted power is limited corollary 1 may not hold and therefore the traffic capacities of the forward and reverse links may not be equal.

Note that if there are only two base stations at most two mobiles can share the same channel. Hence at any feasible assignment there will be a number of channels used by only one mobile and the rest of the channels used by two mobiles. An assignment that requires minimum number of channels minimizes the number of mobiles that use a channel by themselves, or equivalently maximizes the number of mobiles that can share a channel. An optimal assignment corresponds to a maximum matching in an appropriate graph.

Associate with each configuration of mobiles a compatibility graph G=(V,E) created as follows. The nodes of the graph are in one-to-one correspondence with the mobiles. An edge connects two mobiles if they can share the same channel with appropriate base station assignment and power selection. A matching $\mathcal{M}$ of the compatibility graph G is a subset of the edges with the property that no two edges of $\mathcal{M}$ share the same node. Every edge in $\mathcal{M}$ is called matched edge. Maximum matching is a matching that has the maximum possible number of edges. Clearly the set of possible assignments are in one-to-one correspondence with the set of matchings. The next theorem follows readily.

Theorem 2 *The minimum number of channels C_N^f is equal to the number of edges in a maximum matching of the corresponding compatibility graph G (cardinality of the maximum matching) plus the number of nodes that are not incident upon any matched edge. The same result holds for the power control case, too.*

The theorem suggests the following way of computing optimal assignments for any configuration of mobiles.

- Create the compatibility graph G by identifying all possible mobile pairs that can use the same channel (using theorem 1).
- Find a *maximum matching* of the compatibility graph G.
- Allocate the same channel to the mobile pairs that correspond to the edges of the maximum matching. For each such pair make the base station and power assignment.

In accordance with the results we presented in theorem 1 the identification of the pairs of mobiles that can share a channel in the forward direction can be based on relation (9) for the non power control case and on relation (10) for the power control case. In the latter case the optimal power assignment is achieved if the transmitted powers P_{1l}^f, P_{2l}^f for a pair of mobiles that share a channel are chosen to satisfy relation (11). As we mentioned before similar expressions hold for the reverse channel, too. Hence connectivity of the compatibility graph is different in the cases of power control, non power control, forward and reverse channel assignment. Note that the computationally

intensive part of the algorithm is the identification of the maximum matching of the compatibility graph.

Next we apply the algorithm in a simple example. Throughout our numerical studies the threshold T is taken equal to 18 dB, unless otherwise indicated.

Example 1 *Consider a system with two base stations and $N = 10$ mobiles distributed on a line, as shown in figure 4. The optimum reverse channel assignment will be identified.*

By v_i, $i = \{1, 2, \ldots, 10\}$ we denote mobile-i while by BS_j, $j = \{1, 2\}$ we denote base station-j. For the system that is depicted in figure 4(a) we identify first all possible pairs of mobiles that can share a common channel in reverse direction. We create the corresponding compatibility graph represented by its adjacency matrix $C : N \times N$, called the compatibility matrix, where:

$$C[i, j] = \begin{cases} 1, & if \;\; v_i \;\; and \;\; v_j \;\; can\ share\ a\ channel, \\ 0, & otherwise. \end{cases}$$

In our case this matrix has the following form:

$$C = \begin{bmatrix} 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 & 1 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 1 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 1 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 1 & 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \end{bmatrix}$$

The corresponding compatibility graph G is shown in figure 4(b). In order to compute the minimum number of channels C_{10}^r needed to accommodate the mobiles of graph G we must first identify the maximum matching of graph G. For the graph of figure 4(b) a maximum matching is

$$\mathcal{M} = \{[v_2, v_9], [v_3, v_6], [v_4, v_7], [v_5, v_8]\},$$

shown as heavy lines in figure 4(b). We let the pairs of mobiles matched by $\mathcal{M}$ share the same channel. For example v_2 transmits to base station 1 on the same channel that v_9 transmits to base station 2. The cardinality of $\mathcal{M}$ is 4. Therefore the total number C_{10}^r of channels is equal to four plus the number of nodes that are not incident upon any matched edge (v_1 and v_{10}), that is $C_{10}^r = 4 + 2 = 6$. □

Most of the cases of the two way channel assignment problem can be reduced to a maximum matching problem in an appropriate compatibility graph. First we consider the non power control case. Initially we assume that each channel can be used in one direction only (either as forward or as reverse channel) and that each mobile may use any base station for the forward and reverse connection. Then the two way channel assignment problem is reduced to the solution of the forward and reverse channel assignment problems and the total number of channels needed to accommodate all the mobiles in the system is equal to $(C_N^f + C_N^r)$.

Now consider the case that a channel can be used in both directions, if this is preferable, and each mobile may use different base stations for the forward and reverse connection. The corresponding compatibility graph is as follows. Each mobile-i is represented by two nodes in graph G that correspond to the forward and reverse connection of the mobile, respectively. An edge connects two nodes if and only if the corresponding mobiles can share the same channel for their communication in the directions indicated by the two nodes. Then the total number of channels needed to satisfy all the interference constraints in the system is equal to the cardinality of the maximum matching of graph G plus the number of nodes that are not incident upon any matched edge.

Now we turn our attention to the cases that the powers are adjustable parameters. First assume that each channel can be used only in one direction. Then the total number of channels needed to accommodate all the mobiles is equal to $2\tilde{C}_N^f$. This happens because as we have seen in Corollary 1 , if two mobiles can share one forward channel then they can share one reverse channel, too.

Lets now consider the most general case that a channel can be used in both directions. First we assume that a mobile should use the same base station for its communication in both directions. As we conclude from theorem 1 if two mobiles can share a forward channel l_1 then they can also share a reverse channel l_2 and therefore they can form a pair of mobiles that can make use of two channels to communicate in both directions. It can also be shown that if a mobile A uses channel l_1 in forward direction and mobile B uses the same channel in reverse direction then these two mobiles can also make use of one channel l_2 in the opposite directions, i.e. mobile A can transmit to base 1 on l_2 and mobile B can receive from base 2 on l_2. However if any of the above conditions are satisfied then those two mobiles can form a pair that can share common channels for their communication. Then we can create a graph G where as nodes we consider all the mobiles and an edge connects two nodes if and only if the corresponding mobiles can share the same channel, in some way. Following similar reasoning as before we can easily see that the problem of finding the optimal joint channel base assignment is reduced to the identification of the maximum matching of the above created graph G. Actually single nodes in the graph G represent those mobiles for which there is no way to share any channel with any other mobile and they should make use of their own channels in both directions. Denote by C_h the cardinality of the maximum matching of G plus the number of the nodes that are not incident upon any matched edge. Then the total number of channels needed to accommodate all the mobiles of graph G is equal to C_h multiplied by two in order to count for the communication in both directions.

If we eliminate the constraint that a mobile should communicate with the same base station in both directions then a procedure similar to the corresponding case under fixed powers can be followed.

2 Evaluation of traffic capacities

In this section the traffic capacities of the forward and reverse channel, as defined in the introduction, are obtained and compared for the cases of power control and non power control. The mobiles are considered to be randomly and uniformly distributed either on a line or on the plane.

The traffic capacities for the power control case for an arbitrary number of mobiles are evaluated by simulation. The corresponding curves are depicted in figure 7. The first curve (identified by the label power control-uni) corresponds to the case that each channel can be used only in one direction, while the second curve (identified by the label power control-bi) corresponds to the case that a channel can be used in both directions. Figure 8 presents the same quantities but using a threshold of 14 dB. Comparing them to those for the non power control case we verify the considerable improvement on the traffic capacity. Moreover the use of bidirectional channels results in a small extra increase in the corresponding capacity.

3 Ongoing work - Further research

A similar approach to the one we developed in order to identify the joint optimal channel base station and power assignment could be also applied in the general case that we consider an arbitrary number of base stations and mobiles. The same steps could be followed but instead of identifying pairs of mobiles that can share a channel and therefore reducing the problem to the solution of a maximum matching on the appropriate compatibility graph, we must identify all the possible subsets of mobiles that can make use of the same channel. Then we have to select those subsets that correspond to the use of the minimum number of channels in the system, as well as to assign base stations and select powers. As the number of base stations and mobiles increases such a procedure becomes intractable. Therefore some heuristic algorithms that approximate the optimal assignment should be devised based on the steps followed by the optimal algorithm for two base stations. Such algorithms will have great practical value. In [6] we provide a heuristic algorithm that actually does the joint resource allocation in a general network with M base stations and N mobiles and we verify the large capacity improvements that can be achieved through the integration of the channel base station and power assignment.

4 Transmission control in intrasatellite links

The asynchronous transmissions scheduling problem arises in the access layer of the crosslinks of satellite networks. Packet switched satellite networks have been studied extensively, initially as highly survivable communication systems at periods of crisis and more recently as the natural solution for providing globe wide wireless communication services. There are several proposals [2, 5, 8] that involve a large number (60-240), low altitude (400 miles), satellites for providing global coverage and/or survivability. The satellite crosslink distances are up to 3300 nmi and the average link lifetime is approximately 7 min. The primary problem in such a system is to find a resource efficient solution to the multiple access/ multiple resource capacity allocation required among the satellites. In similar systems most protocols resolve time overlaps by allowing contention to occur at the receiving end of links, resulting in wasted transmissions. But due to power limitations and to the fact that propagation delay between neighboring satellites is quite large, on the order of hundreds of packet durations, the above mentioned protocols are not applicable to satellite networks. An efficient scheduling scheme should minimize, if not eliminate wasteful transmissions.

An approach to scheduling transmissions called Pseudo-random Scheduling (PRS) was introduced by Binder et al. [2]. A related protocol called Adaptive Receive Node Scheduling (ARNS) Protocol was introduced by Kosowsky et al. [4]. A key aspect of such an approach is that the synchronization problem is solved locally by providing each satellite a schedule for each of its neighboring satellite receivers. This means that each satellite uses a pseudo-random sequence which dictates when it will listen for packet transmissions from each of its neighboring satellites. Moreover, the times when a satellite is in listen mode are composed of non-overlapping periods, the length of each being the time needed to receive a packet. During each period the satellite is assigned exactly one neighbor to listen to, according to its pseudo-random sequence. Therefore, contention is resolved by the division and assignment of satellite's receive time, such that no two satellites ever try to transmit to the same node at the same time.

During the transmit state, a satellite examines the neighbor's sequences and searches for any receive slots. Since the satellite knows the pseudo-random sequences of its neighbors (which means that it knows the times that each of its neighbors will be listening to it), and can calculate the propagation delay to each of its neighbors, it can easily determine when it has opportunities to send to each of its neighbors. Every node merges the transmission opportunities from all of its links (since each link is synchronized independently) and therefore two or more transmission opportunities may overlap. If no transmit or receive opportunity exists for a satellite, a satellite may use this time to schedule communications with terminals. If a transmit opportunity exists, but a satellite has no traffic in its queue for that neighbor, then this idle time can be used to scan for new satellites or terminals.

In such an environment we need to consider methods for a satellite to schedule packet transmissions to neighboring satellites that eliminate transmission overlaps in time and maximize the throughput. A satellite in transmit mode corresponds to the server of our queueing network model

while the neighbors of the satellite correspond to the different parallel queues. The beginning of a transmission opportunity to a neighbor represents the connectivity time instance of the corresponding queue.

A queueing model suitable for communication networks with asynchronous transmissions is considered. M parallel queues receive service from a single server. The service times are deterministic and equal to τ for all queues. The service of a queue can be initiated only at the connectivity time instances of the queue. The connectivity instances may differ from queue to queue and they are arbitrary in general. The connectivity instances for queue-i are represented by a non-decreasing sequence $\{t_n^i\}_{n=1}^{\infty}$. The aggregation of the above individual sequences of connectivity instances is represented by the sequence $\{(t_n, i_n)\}_{n=1}^{\infty}$ of pairs of random variables, where $\{t_n\}_{n=1}^{\infty}$ is the superposition of the sequences $\{t_n^1\}_{n=1}^{\infty}, \ldots, \{t_n^M\}_{n=1}^{\infty}$ and $i_n, n = 1, 2, \ldots$ is a sequence of M-valued random variables that represent the type of the connectivity times (the queue that may receive service). The service of different packets cannot overlap in time. Therefore scheduling is required to resolve potential contention of transmissions initiated at closely spaced (closer than τ) connectivity instances, and eliminate transmission overlaps in time. A system with three queues is illustrated in figure 9.

A special case of the above model is when the connectivity instances are synchronized to occur only at the beginning of slots of duration τ, or in other words when for every n,k we have $| t_n - t_k | = l\tau$ for some $l = 0, 1, \ldots$. This case has been considered in [10] and [9] and it will be referred to as the synchronous case here. In the synchronous case the services at different slots do not overlap and can be scheduled independently. The issue is how to schedule the server in every slot such that a sufficient fraction of the server capacity is provided at each queue at the slots that the queue may receive service. In [10] a maximum throughput allocation policy was given for the single server system and in [9] a network with changing topology and synchronized connectivity instances was studied. The maximum throughput policy was scheduling the services at every slot based on the connectivity process and the state of the system at that slot only. This is not the case for asynchronous connectivity processes.

Carr and Hajek [3] considered the asynchronous system. They studied several scheduling schemes ranging from simple greedy policies that allocate the server to the first available connectivity instance, to more sophisticated schemes that take the future of the connectivity process under consideration. The throughput of all the different scheduling schemes was evaluated for connectivity processes of Poisson type. It was demonstrated that due to the potential partial overlap of services at different connectivity instances in the asynchronous case, it is possible to improve the throughput compared to what is achievable by strictly non anticipative policies, if the services in longer time intervals are scheduled jointly.

In our work we identified a class of optimal policies, the Anticipative Adaptive Horizon (AAH) policies, for the asynchronous system. Two issues arise in the scheduling of such a system. One is the selection of a sufficiently long scheduling horizon, in order to sustain the traffic load. The other is the allocation of the service capacity to the different queues in a fair manner, in accordance to the

loading of the queues. The AAH policy determines the scheduling horizon and the server allocation adaptively, and it achieves maximum throughput. The transmissions are scheduled in cycles. The cycle length is an increasing function of the system backlog at the beginning of the cycle, that reflects the traffic load. The transmissions are selected such that the weighted throughput in the cycle is maximized, where the weights are equal to the queue lengths in the beginning of the cycle. The scheduling of a cycle is based on the solution of a maximum weighted independent set problem on a colored interval graph that sufficiently represents the connectivity process during the cycle. Unlike the case of general graphs, the computation of the maximum independent set problem on interval graphs can be done in polynomial time and the AAH policy is efficient. The system is studied for Poisson exogenous arrivals. It is shown that the AAH policy maximizes the long term throughput for Poisson and periodic connectivity processes. The performance of the policy and the average scheduling complexity are studied by simulation. It is shown that by adjusting certain parameters of the policy, towards increasing the scheduling horizon, the average delay decreases while the scheduling complexity increases.

5 Scheduling policies and throughput regions

A scheduling policy resolves possible contention of services initiated at connectivity instances which are closer than a packet transmission time. In general scheduling policy is any random subsequence $\{(t_{n_j}, i_{n_j})\}_{j=1}^{\infty}$ of the connectivity process such that $\mid t_{n_{j+1}} - t_{n_j} \mid > \tau$ with probability one, where $t'_{n_j}s$ are the service instances scheduled by the policy. Let $\mathcal{L}$ be the class of all scheduling policies. By $S_i^\pi(t_1, t_2)$ we denote the number of services to queue-i that are scheduled to initiate during the interval $[t_1, t_2 - \tau)$ and therefore finish before t_2, under policy π. Define $S_i^\pi(t) = S_i^\pi(0, t)$. Let also $S^\pi(t_1, t_2)$ and $S^\pi(t)$ be the corresponding service vectors. The effect of the scheduling during an interval (t_1, t_2) on the throughput of the system is completely represented by the corresponding service vector for that interval. The collection of all possible service vectors represents all the different scheduling options for that interval; scheduling amounts to selecting one such vector.

The collection of all feasible service vectors in the interval $[t_1, t_2)$ can be sufficiently represented in terms of the colored interval graph that corresponds to that interval, which is denoted by $G(t_1, t_2)$. This graph contains one node for each connectivity instance in the interval $[t_1, t_2 - \tau)$. The node (t_n, i_n) is *colored* by i_n, the type of the connectivity instance t_n. Two nodes (t_n, i_n) and (t_k, i_k) of the graph are adjacent if and only if the difference of the corresponding connectivity instances is smaller than τ, that is if $\mid t_n - t_k \mid < \tau$. The collection of all service vectors $S^\pi(t_1, t_2)$, $\pi \in \mathcal{L}$ corresponds to the collection of independent sets of the colored interval graph $G(t_1, t_2)$. An independent set of the graph $G(t_1, t_2)$ is any subset of its nodes which contains only pairwise nonadjacent nodes, that is any subset of nodes with no two nodes in the set connected by an edge in $G(t_1, t_2)$. One feasible service vector corresponds to each independent set. The number of type-i nodes of the independent set is equal to the i-th element $S_i^\pi(t_1, t_2)$ of $S^\pi(t_1, t_2)$. The collection of all service vectors associated, in the above sense, with a graph G is denoted by S^G. These entities are illustrated in figure 10.

Note that the graph $G(t_1, t_2)$ depends on the connectivity process and is therefore a random object. Let $H(t_1, t_2)$ be the collection of all possible colored interval graphs that may arise in the interval $[t_1, t_2)$. By convention we denote $G(0, t)$ and $H(0, t)$ by $G(t)$ and $H(t)$ respectively. The probability that the colored interval graph $G(t)$ is equal to $G \in H(t)$ is denoted as

$$p^G(t) = Pr\{G(t) = G\}, \quad G \in H(t).$$

The above probability distribution is implied by the statistics of the connectivity process.

5.1 System throughput region

Assume that an actual packet is served at each time instant scheduled by the policy such that the server finds no empty queue. The expected potential throughput vector during the interval $[t_1, t_2)$ under policy $\pi \in \mathcal{L}$ is defined as

$$\lambda^\pi(t_1, t_2) = \frac{1}{t_2 - t_1} E[S^\pi(t_1, t_2)].$$

In continual operation, when packets continuously arrive in the system and the queues may be empty when they are scheduled for service, $\lambda^\pi(t_1, t_2)$ may be viewed as an upper bound to the achievable throughput. By convention we denote $\lambda^\pi(0, t)$ by $\lambda^\pi(t)$. The region of the achievable throughput vectors in the interval $[t_1, t_2)$ is defined as

$$\Lambda(t_1, t_2) = \{\lambda^\pi(t_1, t_2) \ : \ \pi \in \mathcal{L}\}$$

and we denote $\Lambda(0, t)$ by $\Lambda(t)$.

A useful representation of the achievable throughput vectors in the interval $[0, t)$ is provided by the following theorem.

Theorem 3 *A throughput vector λ belongs to the region $\Lambda(t)$ if and only if there exist vectors λ^G in the convex hull $co(S^G)$ of S^G, for all $G \in H(t)$, such that*

$$\lambda = \frac{1}{t} \sum_{G \in H(t)} \lambda^G p^G(t). \tag{12}$$

The system throughput region contains all the throughput vectors achievable during a long run operation of the system. It is defined as

$$\Lambda = \limsup_{t \to \infty} \Lambda(t).$$

Explicit characterization of Λ is not always possible and depends on the statistics of the connectivity processes. The difficulty in obtaining such characterization is due to the fact that, in general, there might be indefinitely long sequences of connectivity instances which are interdependent in the sense that services at successive instances overlap. An important special case of connectivity processes for which the characterization of Λ is fairly simple is for deterministic periodic processes. Consider a connectivity process for which there is an integer $k > 0$ and $T > 0$ such that for every n

$$t_{n+k} = t_n + T, \quad i_{n+k} = i_n \quad and$$

$$t_{lk+1} - t_{lk} > \tau \quad l = 1, 2, \ldots$$

It is not difficult to show that in this case the system throughput region coincides with the region of throughput vectors achievable in one period.

Due to the asynchronous nature of the connectivity instances the service schedules in consecutive intervals $[t_1, t_2)$ and $[t_2, t_3)$ may be interdependent since the last service scheduled in the first interval may conflict with the first service scheduled in the second interval. Considering the whole interval $[t_1, t_3)$ for scheduling instead of the intervals $[t_1, t_2)$ and $[t_2, t_3)$ disjointly will lead to more feasible options in allocating the server, consequently increasing the throughput. In other words the collection $S^{G(t_1,t_3)}$ of all service vectors in the interval $[t_1, t_3)$ is strictly larger in general than the collection of all service vectors which are sums of any two service vectors from $S^{G(t_1,t_2)}$ and $S^{G(t_2,t_3)}$ respectively. In general the longer future time horizon we are considering in scheduling services at the present, the larger the throughput that can be achieved. When the system is heavily loaded, that is when the throughput vector is close to the boundary of Λ, then long time horizons should be considered for scheduling. After the appropriate horizon is selected then there is the issue of how to select the service vector such that sufficient fractions of the service capacity are allocated to the individual queues. Unless there is an explicit characterization of the throughput region, it is not possible to determine either how heavily loaded the system is or how to achieve the desirable allocation fractions. In the following we present an adaptive policy that resolves both issues and achieves maximum throughput.

6 Anticipative policies with adaptive scheduling horizon

We consider the system with exogenous arrivals where packets arrive at each queue-i according to a Poisson process of rate λ_i. Without loss of generality, assume $\lambda_i > 0, i = 1, \ldots, M$. The connectivity instances at each queue are assumed to occur with rate μ_i. The adaptive scheduling policies operate in cycles and the scheduling in each cycle is done independently. The k-th cycle lasts from time τ_k to $\tau_{k+1} = \tau_k + h_k$, where by h_k we denote the scheduling horizon in the k-th cycle. The parameter h_k changes from cycle to cycle. The length of the scheduling interval is determined at the beginning of the interval based on the length of the queues at that time. Denote by $X = \{X(t),\ t \geq 0\}$ the queue length process, where $X(t) = (X_1(t), \ldots, X_M(t))$ and $X_i(t)$ is the number of packets of queue-i at time t. The horizon h_k is $h_k = g(X(\tau_k))$, where $g : Z_+^M \to R^+$. The

function g distinguishes the different scheduling policies. It should satisfy the following properties for the policy to have the desirable throughput properties:

$$\lim_{\|X(t)\|\to\infty} \frac{g(X(t))}{\|X(t)\|} = 0, \qquad \lim_{\|X(t)\|\to\infty} g(X(t)) = \infty \tag{13}$$

To avoid trivialities we also assume that $g(x) > 0$ for $x \in Z_+^M$. Within a scheduling interval $[\tau_k, \tau_{k+1})$, the service vector $S(k)$ selected by the policy is such that

$$S(k) = \arg \max_{S \in S^{G(\tau_k, \tau_{k+1})}} \{X(\tau_k)S\}. \tag{14}$$

This means that the service vector $S(k)$ is selected such that the weighted throughput of the system in the corresponding scheduling interval is maximized, where the weights of the services of each queue are equal to the individual queue lengths at the beginning of the scheduling interval. Note that a service vector $S(k)$ is not always realizable since it is possible that the number of services $S_i(k)$ for queue i is larger than the number of packets that will be available at that queue during the interval $[\tau_k, \tau_{k+1})$. If we denote by $R_i(k)$ the number of actual services provided in the interval $[\tau_k, \tau_{k+1})$ then we have

$$R(k) \geq \min\{S(k), X(\tau_k)\} \tag{15}$$

where $R(k) = (R_1(k), .., R_M(k))$ and the min is applied componentwise. Some complexity issues regarding the computation the service vector are discussed in section 9. The number of packets at queue-i evolves with time according to the following equation

$$X_i(\tau_k) = X_i(\tau_{k-1}) - R_i(k) + A_i(k), \tag{16}$$

where by $A_i(k)$ we denote the number of packets that arrived in queue i during the kth cycle. For Poisson and deterministic periodic connectivity processes the AAH policy achieves maximum throughput in the sense that guarantees stability of the system for any arrival rate vector in the interior of Λ.

7 Poisson connectivity processes

Assume that the process of connectivity time instances is Poisson and independent of the process of indicator variables that indicate the type of the connectivities; the latter process is assumed to be i.i.d. The AAH policy achieves maximum throughput in the following sense.

Theorem 4 *If an arrival vector λ belongs to the interior of the system throughput region Λ, then under policy AAH the queue length process converges in distribution to a random vector $\tilde{X}$ such that*

$$E\tilde{X} < \infty.$$

□

The proof of the theorem is preceded by some preliminary results. We first consider the process $\hat{X} = \{X(\tau_k)\}_{k=1}^{\infty}$, where $X(\tau_k) = (X_1(\tau_k), \ldots, X_M(\tau_k))$, of the queue length vectors observed at the beginning of the scheduling cycles and prove some results regarding the steady state moments of $\hat{X}$, as stated in the following Theorem 5. Then we turn our attention to the process $\{X(t),\ t \geq 0\}$ and using the regenerative approach ([1]), we conclude Theorem 4.

Theorem 5 *The process $\{X(\tau_k)\}_{k=1}^{\infty}$ is an irreducible aperiodic positive recurrent Markov chain. Under the stationary distribution it holds*

$$E[\|X(\tau_k)\| g(X(\tau_k))] < \infty. \tag{17}$$

□

8 Periodic connectivity processes

Consider a periodic connectivity process, as it has been defined in the end of section 5. There is an integer $k > 0$ and $T > 0$ such that for every n

$$t_{n+k} = t_n + T, \quad i_{n+k} = i_n \quad and$$

$$t_{lk+1} - t_{lk} > \tau \quad l = 1, 2, \ldots$$

Recall that in the end of section 5 it was shown that the throughput region of the system with periodic connectivities is $\Lambda = \frac{1}{T} co(S^{G(T)})$. Under the AAH policy the system is stable as stated in the following theorem.

Theorem 6 *If an arrival vector λ belongs to the interior of the system throughput region $\frac{1}{T} co(S^{G(T)})$, then under policy AAH the queue length process $\{X(t_i)\}_{i=0}^{\infty}$ converges in distribution to a random vector $\tilde{X}$ such that*

$$E\tilde{X} < \infty.$$

□

The proof of the theorem follows the same steps as that of theorem 4 and it will just be outlined here. Let $\{\hat{\tau}_k\}_{k=1}^{\infty}$ be a subsequence of the connectivity instances, where $\hat{\tau}_k$ is the connectivity instance closest to the beginning of the kth cycle. Notice that the process $\{X(\hat{\tau}_k)\}_{k=1}^{\infty}$ is a homogeneous aperiodic Markov chain. Using basically the same approach used in the proof of theorem 5, we can show that this Markov chain is positive recurrent and has finite first moment under the stationary distribution. The proof is concluded using the regenerative nature of $\{X(t_i)\}_{i=1}^{\infty}$.

If the arrival rate vector belongs to the interior of $\frac{1}{T}co(S^{G(T)})$ then it can be expressed as a convex combination of vectors in $S^{G(T)}$,

$$\lambda = \sum_{e \in S^{G(T)}} a_e^G e, \quad a_e^G \geq 0, \quad and \quad \sum_{e \in S^{G(T)}} a_e^G \leq 1.$$

It is not hard to show that the randomized policy that schedules each period independently by selecting vector e with probability a_e^G stabilizes the network as well. The AAH achieves the same result without needing to know the arrival rate vector. Notice that an explicit characterization of the throughput region as above, does not exist for Poisson connectivities.

One question that remains to be addressed is the magnitude of improvement on the achievable throughput by the use of anticipative policies. Carr and Hajek [3] report simulation results for two queues where it is shown that the throughput region achieved by the adaptive threshold policy, defined in their paper, is very close to the maximum throughput region of the system. Furthermore it is mentioned that similar behavior has been observed in systems with larger number of queues.

If the connectivity process is not Poisson then considerable improvements may be observed by using the AAH policy over what can be achieved by using a non anticipative policy. In the following we demonstrate this point by presenting an example where the AAH policy increases the throughput of a queue by 66 percent over what is achievable by the best threshold policy. Consider a system with three queues and periodic connectivity process, of period T. We denote by u the common overlapping time of two successive overlapped opportunities for all opportunity periods. It is not hard to verify that the throughput vector $(\lambda_1, \lambda_2, \lambda_3) = (\lambda, \lambda, \lambda)$ is within the throughput region of the system, and therefore achievable by the AAH policy, for every $\lambda < \frac{1}{T}$ in packets per time units. In the following we argue that if

$$\frac{0.97}{T} < \lambda_1, \lambda_3 < \frac{1}{T} \tag{18}$$

then the system becomes unstable under any threshold policy whenever $\lambda_2 > \frac{0.6}{T}$. Therefore the throughput of queue 2 cannot be larger than $\frac{0.6}{T}$ compared to $\frac{1}{T}$ that is the tight upper bound on the throughput of queue 2 under the AAH policy, for the same range of λ_1, λ_3.

A threshold policy, as defined in [3], is specified by a set of of thresholds $T_{ij},\ i = 1, \ldots, M,\ j = 1, \ldots, M$. At a connectivity instant t of queue j, service is provided to that queue if the system is idle or if the queue i that is under service at t has received service for an amount of time less than T_{ij}. Note that for the connectivity process under consideration it only matters for the operation of the policy whether a threshold is larger or smaller than $\tau - u$ and not its actual value. Four different operation modes may appear depending on the values of the thresholds.

A. Queue 1 is not preempted by queue 2 and queue2 is not preempted by queue 3.

B. Queue 2 preempts queue 1 and queue 2 is not preempted by queue 3.

C. Queue 1 preempts queue 2 and queue 3 preempts queue 2.

D. Queue 1 is preempted by queue 2 and queue 2 is preempted by queue 3.

In cases A,B and C we argue briefly in the following about our claim. In case D we verified our claim by simulation.

In cases A and B the existence of queue 3 does not affect the operation of queues 1 and 2. Furthermore because of the symmetry, the utilization of each opportunity of queue 2 is $\frac{T\lambda_2}{2}$ when the system is stable. Since queue 3 is served only whenever the preceding queue 2 opportunity is idle, the system is stable if $\frac{\lambda_2}{2} + \lambda_3 < \frac{1}{T}$, that is $\lambda_2 < 2(\frac{1}{T} - \lambda_3)$. If $\frac{0.97}{T} < \lambda_3 < \frac{1}{T}$ clearly $\lambda_2 < \frac{0.6}{T}$ for stability.

In case C queues 1 and 3 operate as if queue 2 were absent. The utilization of each queue 1 opportunity is $\frac{T\lambda_1}{2}$, while the utilization of each queue 3 opportunity is clearly $T\lambda_3$. The total throughput of queue 2 is $\lambda_2^1 + \lambda_2^2$, where λ_2^1 is due to the packets transmitted at the queue 2 opportunity that overlaps with queue 3 opportunity, while λ_2^2 is due to the packets transmitted at the other opportunity of queue 2. Clearly for stability we need $\lambda_2^1 < \frac{1}{T} - \lambda_3$ and $\lambda_2^2 < \frac{1}{T} - \frac{\lambda_1}{2}$, that is $\lambda_2 < \frac{2}{T} - (\lambda_3 + \frac{\lambda_1}{2})$ and for λ_1, λ_3 that satisfy (18), $\lambda_2 < \frac{0.6}{T}$.

9 Scheduling complexity of AAH and modifications

An important consideration regarding the AAH policy is the computational complexity for its implementation. In every cycle of the policy, the maximum weighted service vector need to be evaluated in relation (14). Notice that evaluating this maximum in the kth cycle, is equivalent to solving the maximum weighted independent set problem in the colored interval graph $G(\tau_k, \tau_{k+1})$ where the weight of a node is equal to the queue length of the queue that corresponds to the node. The component $S_i(k)$ of the corresponding service vector is the number of nodes associated with queue i in the maximum weighted independent set. Unlike the maximum weighted independent set problem for general graphs which is NP-complete, for interval graphs the problem can be solved in polynomial time and more specifically with an algorithm of complexity $O(N^2)$ where N is the number of nodes of the graph, that is the number of connectivity instances in the interval $[\tau_k, \tau_{k+1}]$. In fact Carr and Hajek, in the context of a class of policies proposed in [3], provide a dynamic programming type of algorithm that solves this problem. We use that algorithm in the implementation of the policy that we used in the simulations.

¿From the fact that the complexity of scheduling a cycle grows with the square of the number of connectivity points in the cycle, therefore with the square of the cycle length, we may deduce that the scheduling complexity per time unit grows as the cycle length increases. Recalling by the definition of the policy that the cycle length increases as the load increases, we may deduce that the computational complexity of the policy increases with the load. It turns out though that the computational complexity per time unit of the AAH policy is upper bounded by a constant, both for Poisson and periodic connectivity processes, as long as the arrival rate vector lies in the interior of Λ.

The key observation is that if the graph $G(\tau_k, \tau_{k+1})$ consists of several disjoint components then the maximum weighted independent set can be computed separately for each component graph and the union of those independent sets is the maximum weighted independent set of the graph. The scheduling complexity of a cycle is of the order of the sum of the squares of the nodes of the different connected component graphs.

Consider periodic connectivity processes with period equal to T and K connectivity instances per period. In this case the maximum number of nodes in a connected component of the colored interval graph of any cycle is equal to K. The scheduling complexity per time unit is of the order of $O(K^2/T)$ and it is independent of the load.

In the case of Poisson connectivity processes, consider maximal sequences of connectivity instances in which adjacent connectivity instances are closer than a packet length. Consider the subsequence of connectivity times specified by the sequence of indices

$$m_1 = 1,\ m_k = \min\{i : t_i > t_{m_{k-1}},\ t_i - t_{i-1} > \tau\}.$$

The sequence of connectivity times $t_{m_k}, t_{m_{k+1}}, .., t_{m_{k+1}-1}$ can be scheduled independently of the rest connectivity instances, without any degradation of the throughput. In other words the size of a connected component of the colored interval graph of any cycle is stochastically smaller than the number of connectivities in a sequence as the above. A sequence of connectivities as above will be called Conflict Resolution Period (CRP) in the following. Hence Let R be the number of connectivity instances in such a run and T the corresponding time length. The scheduling complexity per time unit then is upper bounded by $E[R/T]$. Note that this bound is independent of the load and is determined from the connectivity process.

The following modified version of the AAH policy takes under consideration the fact that the scheduling of different conflict resolution periods can be done independently. The modified policy is similar to AAH except of the determination of the cycle length that is done as follows. Let t_i be the earliest connectivity instant after time τ_k with the property $t_{i+1} - t_i > \tau$. Then the length of the kth cycle is selected to be $h_k = \min\{t_i - \tau_k, g(X(\tau_k))\}$. The modified policy will be called MAAH in the following. The MAAH policy has superior performance over the AAH, as shown in the simulation study.

10 Simulation Results

The queueing delay and the scheduling complexity were studied by simulation. Both the AAH and the MAAH policies were considered. The scheduling horizon was selected by functions of the following type: $g(\mathbf{x}) = (\sum_{i=1}^{M} x_i)^\alpha$, $0 < \alpha < 1$. The parameter α determines the length of the scheduling horizon. Its effect on the queueing delay and the throughput was evaluated. A symmetric system with three queues was studied. The packet length was taken equal to the time

unit. Poisson connectivities were considered, with rate $\mu = 1.2$ and uniformly distributed for the different queues.

In figure 11 we see plots of the average queue length as a function of the total load for the AAH and MAAH policies respectively. Note that by Little's Law we can readily deduce the corresponding delay plots which have the same qualitative behavior with the average queue length plots. The three queues are equally loaded. The average queue length is plotted for several different values of the parameter α. Clearly the policy MAAH outperforms the policy AAH. Furthermore the performance is improved as the parameter α increases. Recall that the difference between MAAH and AAH is that while MAAH never schedules more than one CRP at a time, the AAH may schedule several CRP's at a time based on the queue lengths at the beginning of the scheduling period. The MAAH schedules each CRP separately using the backlog information at the beginning of the CRP. Hence MAAH uses more updated state information than AAH for scheduling.

11 Ongoing work- further research

The AAH policies described above achieve maximum throughput, stabilizing the system for all stabilizable traffic loads. Furthermore they are adaptive and they do not rely on knowledge of the traffic parameters. The scheduling complexity of the policies increases with the load. This is inevitable in asynchronous connectivity systems, since an increase of the load necessitates the joint scheduling of opportunities in longer scheduling periods. Nevertheless the scheduling complexity is bounded by a constant independent of the arrival rates. AAH (and MAAH) are parametrized class of policies. Even though the maximum throughput property holds for the whole range of variation of the parameter α, the average delay as well as the scheduling complexity varies with the parameter α for a fixed load. By selecting the parameter α, as we did in the simulations, we may achieve any desirable trade-off between the scheduling complexity and delay.

The maximum throughput property has been shown here for two types of connectivity processes, Poisson and deterministic periodic processes. We believe that AAH retains the maximum throughput property for a large class of connectivity processes including renewal connectivities or connectivity process with dependent inter-connectivity times. Another important issue to be investigated is the selection of the connectivity process. By that we refer to both the selection of the statistics of the connectivity time instances as well as of the fractions that correspond to each queue. It is expected that the connectivity process itself will have a significant impact on the performance of the system, in addition to the scheduling policy.

References

[1] S. Asmussen. *Applied Probability and Queues.* John Wiley and Sons, 1987.

[2] R. Binder, S.D. Huffman, I. Gurantz, and P.R. Vena. Crosslink architectures for a multiple satellite system. *Proceeding of the IEEE*, 75:74–81, January 1987.

[3] M. Carr and B. Hajek. Scheduling with asynchronous service opportunities with applications to multiple satellite systems. *IEEE Transactions on Automatic Control*, 38(12):1820–1833, December 1993.

[4] R.P Kosowsky, I.M. Jacobs, and K.S. Gilhousen. Arns: A new link layer protocol. pages 515–519. In IEEE MILCOM-88, 1988.

[5] R. J. Leopold. Low-earth orbit global cellular communications network. In *ICC'91*, pages 1098–1102, Denver, Colorado, June 1991.

[6] S. Papavasiliou and L. Tassiulas. Improving the capacity in wireless networks through integrated channel, base station and power assignment. *IEEE Transactions on Vehicular Technology*, 1996. (submitted).

[7] S. Papavasiliou and L. Tassiulas. Joint optimal channel allocation power control and base station assignement in wireless networks. *IEEE/ACM Transactions on Networking*, October 1996. (to appear).

[8] D. E. Sterling and J. E. Hatlelid. The iridium system - a revolutionary satellite communications system developed with innovative applications of technology. pages 436–440. MILCOM'91, 1991.

[9] L. Tassiulas. Scheduling and performance limits of networks with constantly changing topology. *IEEE Transactions on Information Theory*, 1996. to appear.

[10] L. Tassiulas and A. Ephremides. Dynamic server allocation to parallel queues with randomly varying connectivity. *IEEE Trans. on Information Theory*, 39(2):466–478, March 1993.

[11] L. Tassiulas and S. Papavassiliou. Optimal anticipative scheduling with asynchronous transmission opportunities. *IEEE Transactions on Automatic Control*, 40(12):2052–2062, December 1995.

[12] J. Zander. Performance of optimum transmitter power control in cellular radio systems. *IEEE Trans. on Veh. Technol.*, 41(1):57–62, February 1992.

Personnel supported

- **Faculty:** Leandros Tassiulas
- **Students (PhD) :** Jeong Ryoo, Chi-Jiun Su, Semyon Papavassiliou. (Partially supported).

Publications Acknowledging Grant F49620-95-1-0061

- L. Tassiulas, " Adaptive Back-Pressure Congestion Control Based on Local Information," *IEEE Transactions on Automatic Control*, Vol. 40, No. 2, pp. 236-250, February 1995.
- L. Georgiadis, W. Szpankowski, L. Tassiulas, "A scheduling policy with maximal stability region for ring networks with spatial reuse," *Queueing Systems: Theory and Applications*, Vol. 19, pp. 131-148, 1995.
- L. Tassiulas, J. Joung, "Performance measures and scheduling policies in ring networks," *IEEE/ACM Transactions on Networking* , 1995 (to appear).
- L. Tassiulas, A. Ephremides, "Throughput Properties of a Queueing Network with Distributed Dynamic Routing and Flow Control," *Advances of Applied Probability*, March 1996 (to appear).
- L. Tassiulas, S. Papavassiliou, "Optimal Anticipative Scheduling with Asynchronous Transmission Opportunities," accepted in the *IEEE Transactions on Automatic Control*, 1995 (regular paper).
- L. Tassiulas, L. Georgiadis, "Any work-conserving policy stabilizes the ring with spatial reuse," *IEEE/ACM Transactions on Networking* , 1995 (to appear).
- L. Tassiulas, " Adaptive Routing on the Plane," *Operations Research*, 1995 (to appear).
- L. Tassiulas, "Push-Forward Link-Level Scheduling for Network-wide Performance," accepted in the *IEEE/ACM Transactions on Networking* , 1995.
- L. Georgiadis, W. Szpankowski, L. Tassiulas, "Stability Analysis of Quota Allocation Access Protocols in Ring Networks With Spatial Reuse," accepted in the *IEEE Transactions on Information Theory*, 1995.
- L. Tassiulas, "Scheduling and Performance Limits of Networks with Constantly Changing Topology," Polytechnic Univ. Tech. Rep. CATT-93-62, submitted to the *IEEE Transactions on Information Theory*, 1994.
- S. Papavassiliou, L. Tassiulas, "Joint Optimal Channel Allocation Power Control and Base Station Assignment in Wireless Networks," Polytechnic Univ. TR CATT-93-64, submitted to the *IEEE/ACM Transactions on Networking*, 1994.

- S. Papavasiliou, L. Tassiulas, P. Tandon "Meeting QOS Requirements in a Cellular Network with Reuse Partitioning," *Proceedings of the INFOCOM94*, Toronto, Canada, 1994.

- L. Tassiulas, "Low Complexity Randomized Policies for Maximum Throughput in Queueing Systems with Constraints," (invited) *Proceedings of 32nd Annual Allerton Conference on Communication, Control and Computing*, Urbana Illinois, 1994.

- L. Tassiulas, S. Papavassiliou, "A Dynamic Scheduling Problem in Packet Switched Satellite Networks," in *Proceedings of 33rd CDC*, Lake Buena Vista, Florida, 1994.

- L. Tassiulas, S. Papavassiliou, "The Joint Resource Allocation Problem in Wireless Networks," in *Proceedings of ICC'95*, pp 395-399, Sheattle, Washington, 1995.

- L. Tassiulas, S. Papavassiliou, "Improving the Capacity in Wireless Networks through Integrated Channel Base Station and Power Assignment," in *Proceedings of GLOBECOM' 95*, Singapore, 1995.

Interactions/Transitions

32nd Annual Allerton Conference on Communication, Control and Computing, Urbana Illinois, 1994. Presentation of the paper: *Low Complexity Randomized Policies for Maximum Throughput in Queueing Systems with Constraints*

33rd Conference on Decision and Control, Lake Buena Vista, Florida, 1994. Presentation of the paper: *A Dynamic Scheduling Problem in Packet Switched Satellite Networks*

International Communications Conference 95, Sheattle, Washington, 1995. Presentation of the paper: *The Joint Resource Allocation Problem in Wireless Networks*

Air Force Office of Scientific Research Workshop on *Communications Networks for Command and Control: Signaling/Computing Technologies*, Raleigh, North Carolina, March 1995.

Honors/Awards

IEEE Conference on Computer Communications, INFOCOM '94 Prize Paper award for the paper "Meeting QOS Requirements in a Cellular Network with Reuse Partitioning" coauthored with S. Papavassiliou and P. Tandon.

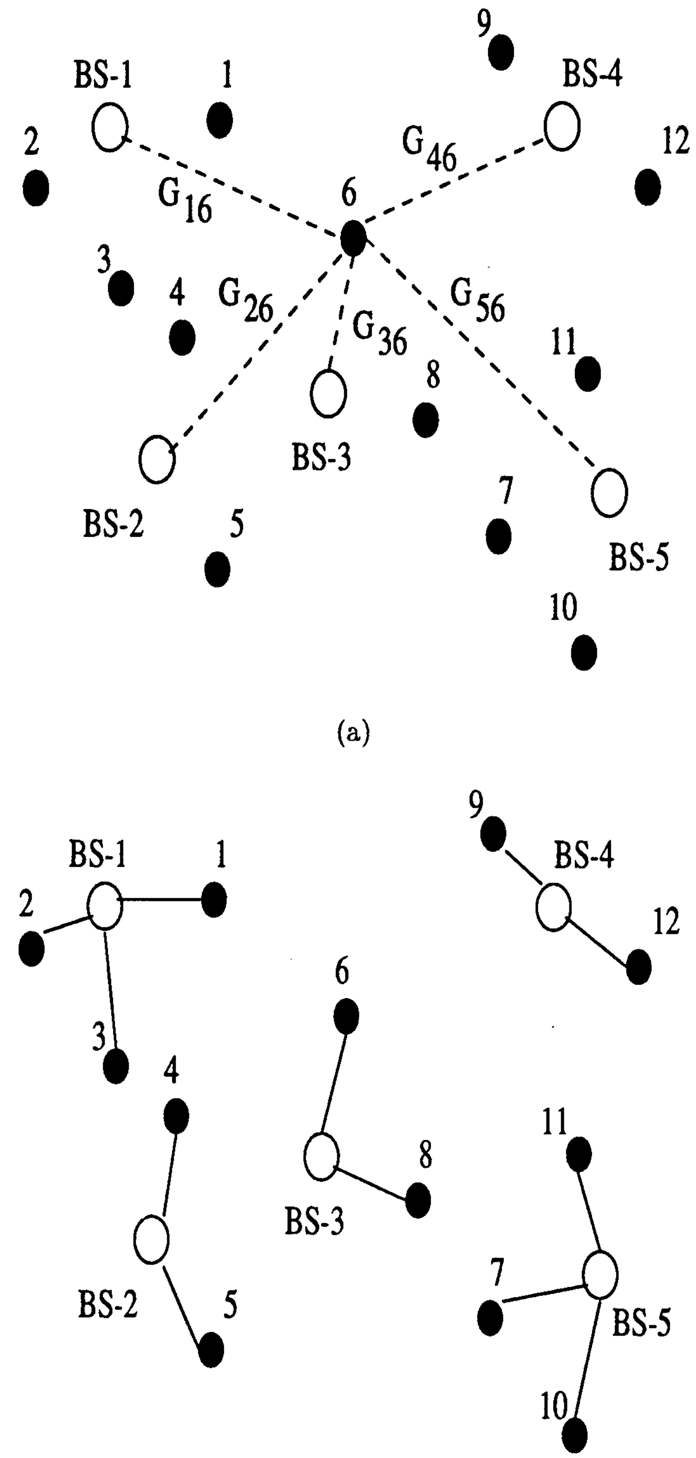


Figure 1: (a) A number of mobiles need to establish a forward and reverse connection with some base stations. The transmission gains G_{ij} are given between any two locations i and j. (b) A base station is selected by each mobile for its forward and its reverse links. The two base stations need not be the same.

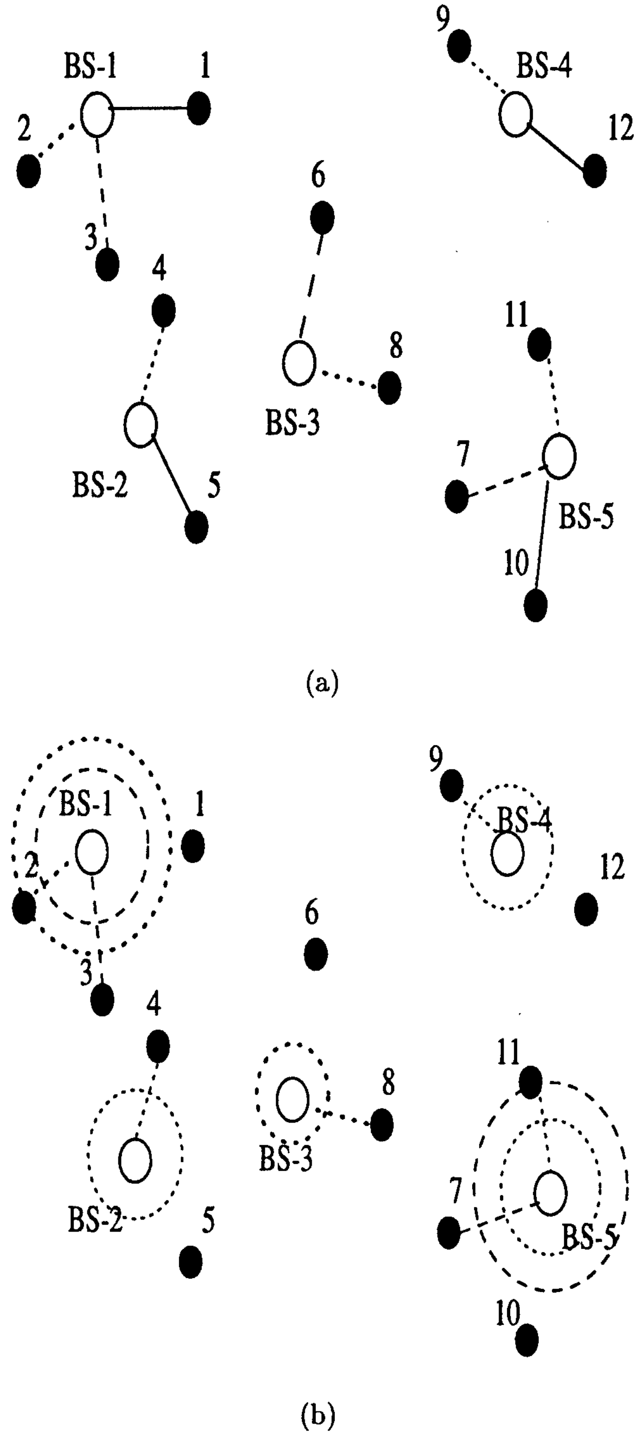


Figure 2: (a) A channel is selected by each link. The cochannel links are represented by the same type of line. (b) The transmission powers of the cochannel links should be selected such that the interference constraints at each mobile are satisfied.

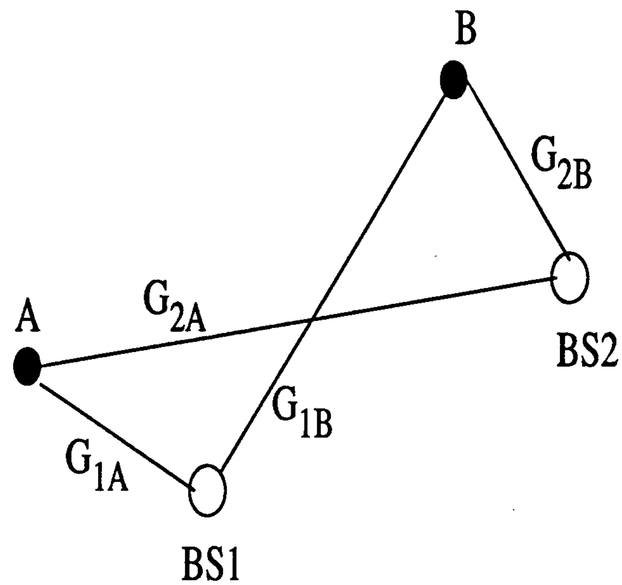


Figure 3: System geometry and link gains.

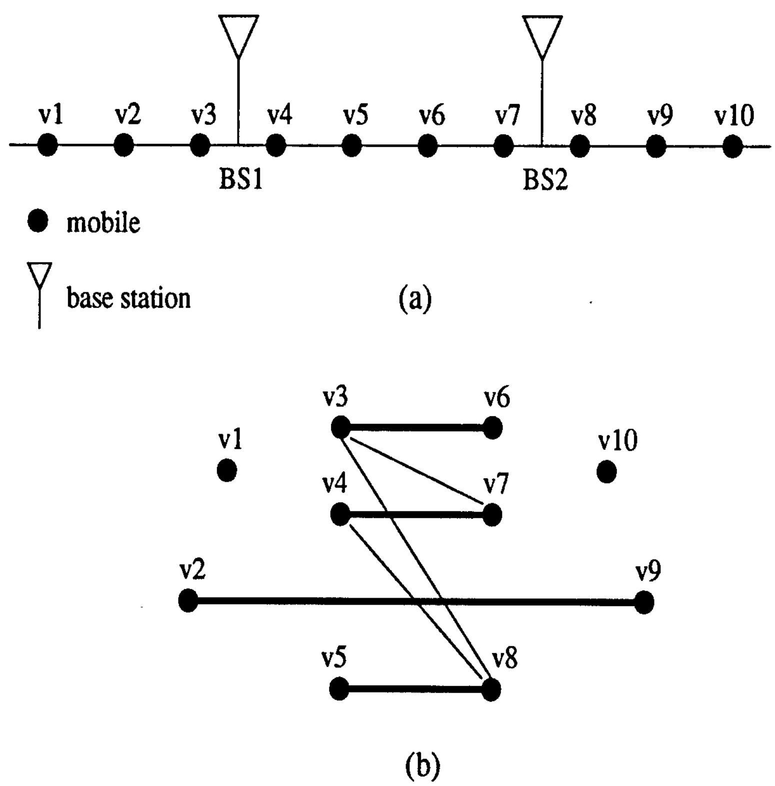


Figure 4: (a) Two base stations and $N = 10$ mobiles distributed on a line. (b) Compatibility graph G - The heavy lines represent a maximum matching.

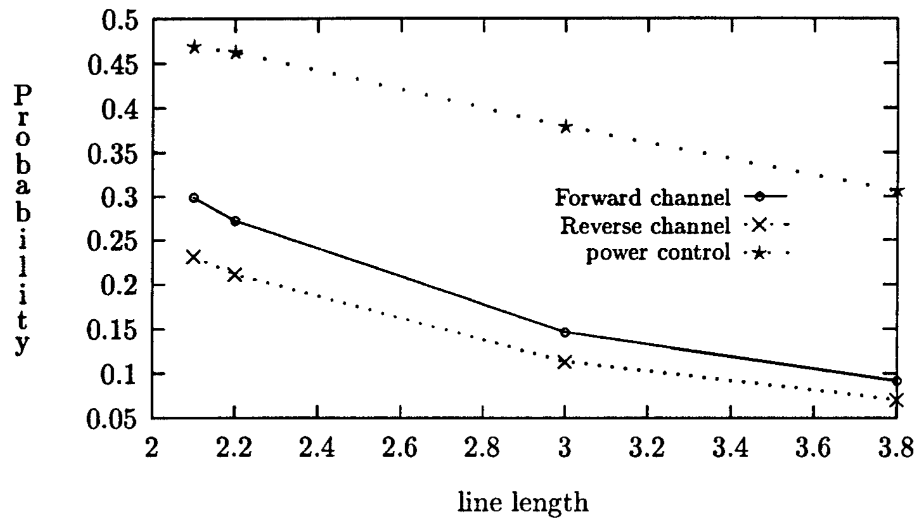


Figure 5: Probability of using only one channel for the communication of two mobiles in one direction (linear case, d=1).

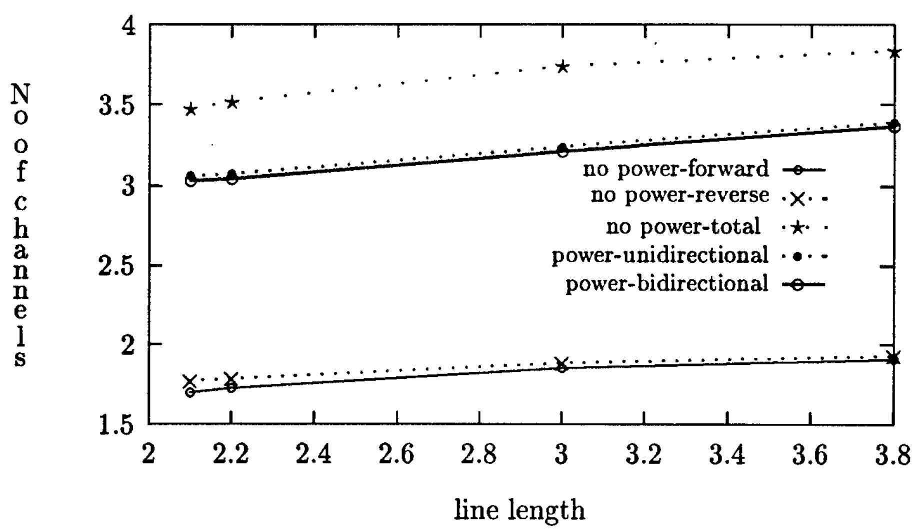


Figure 6: Expected number of channels for the communication of two mobiles (linear case, d=1).

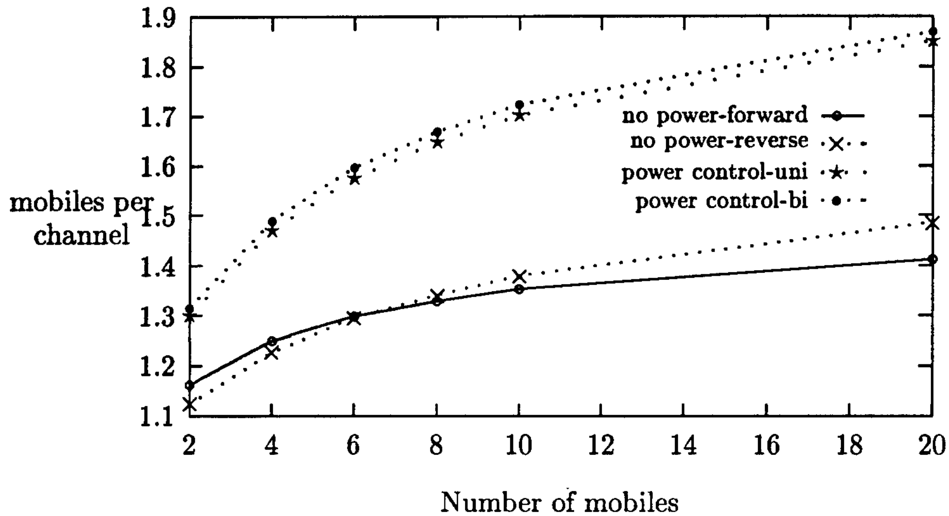


Figure 7: Expected number of mobiles per channel (linear case, d=1) - threshold of 18 dB.

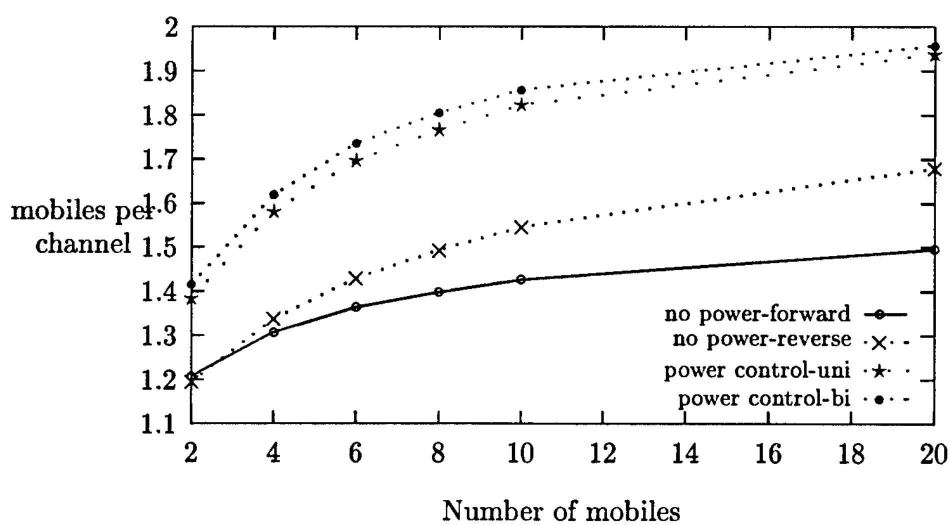


Figure 8: Expected number of mobiles per channel (linear case, d=1) - threshold of 14 dB.

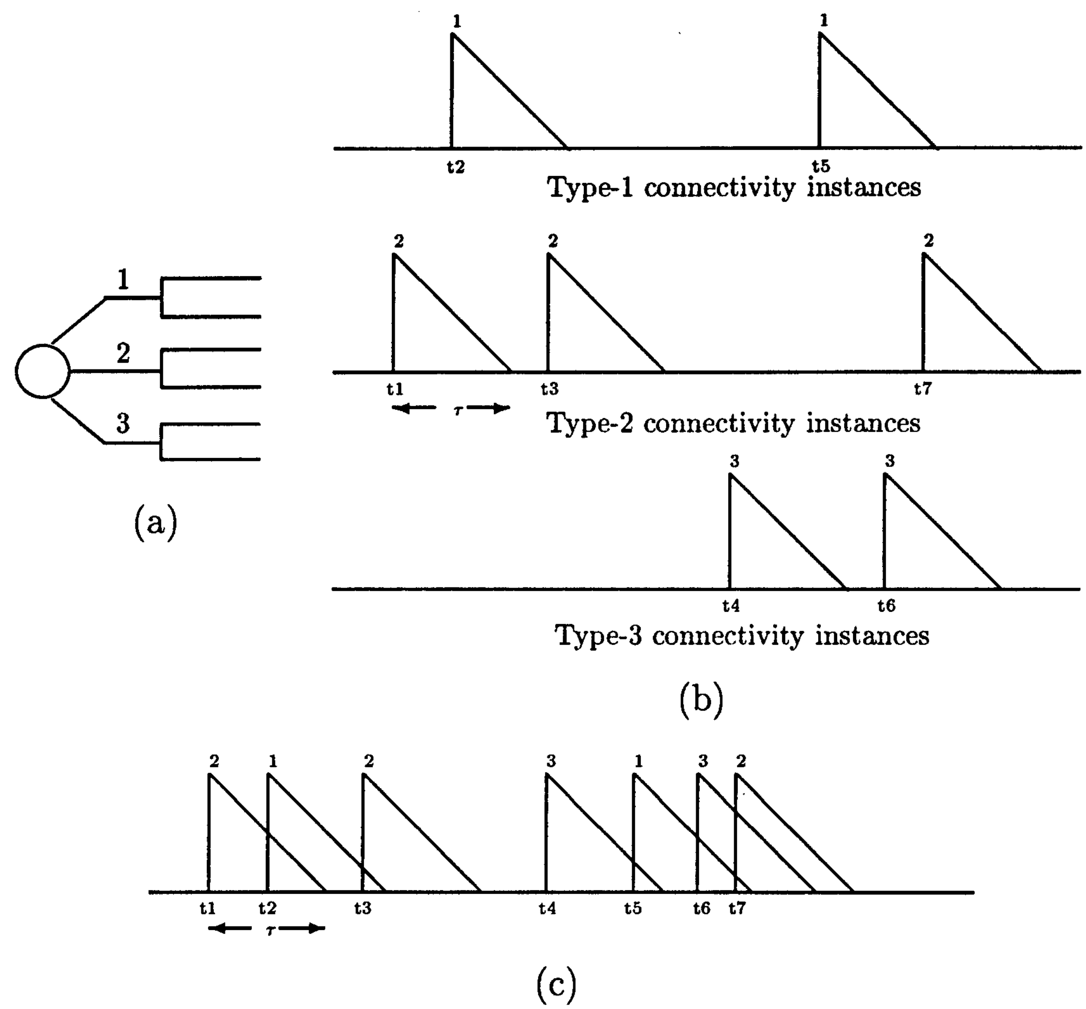
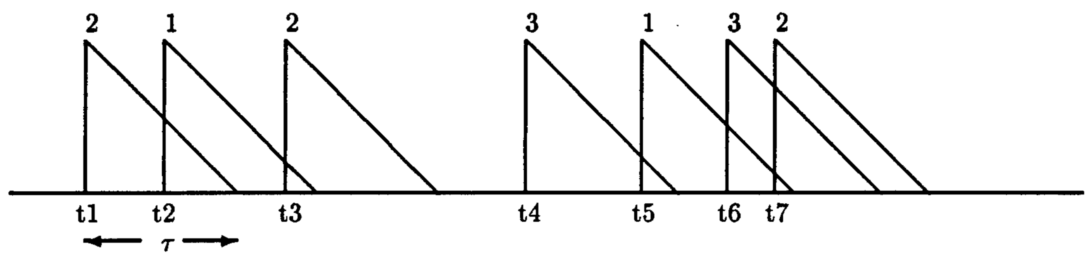


Figure 9: (a) A single server with three queues. (b) The connectivity instances for each queue. Each possible service is represented by a triangular pulse of width τ, where the height at any given time is the residual time for that packet. (c) Aggregate connectivity process.



(a)

(b)

{(t1,2),(t3,2),(t4,3),(t6,3)}	S1=(0,2,2)
{(t1,2),(t3,2),(t4,3),(t7,2)}	S2=(0,3,1)
{(t1,2),(t3,2),(t5,1)}	S3=(1,2,0)
{(t2,1),(t4,3),(t6,3)}	S4=(1,0,2)
{(t2,1),(t5,1)}	S5=(2,0,0)

(c)

Figure 10: (a) Aggregate connectivity process for M=3 queues. (b) The corresponding colored interval graph. (c) Some independent sets of the colored interval graph and the corresponding service vectors.

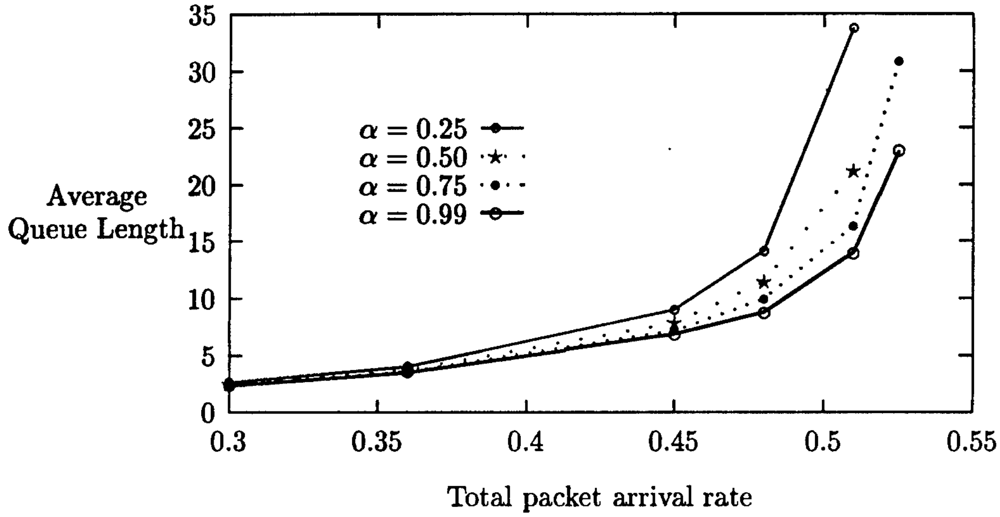


Figure 11: Average total queue length for a system with M=3 queues and $\tau = 1$ under AAH policy. The total opportunity arrival rate is $\mu = 1.2$. All queues are equally loaded.

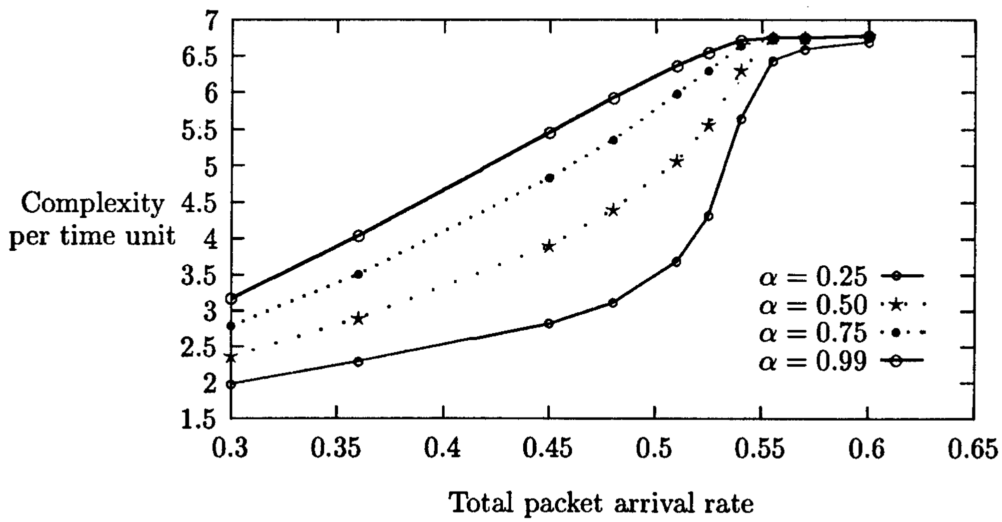


Figure 12: Computation complexity vs. packet arrival rate for a system with M=3 queues and $\tau = 1$. The total opportunity arrival rate is $\mu = 1.2$. All queues are equally loaded.